Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome

Andreas C. W. Jenke*†, Isa M. Stehle*†, Frank Herrmann‡, Tobias Eisenberger*, Armin Baiker*, Jürgen Bode§, Frank O. Fackelmayer‡, and Hans J. Lipps*¶

*Institute of Cell Biology, Witten/Herdecke University, 58448 Witten, Germany; ‡Department of Molecular Cell Biology, Heinrich Pette Institute, 20251 Hamburg, Germany; and §Gesellschaft für Biotechnologische Forschung, D-38124 Braunschweig, Germany

Edited by Mark T. Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved June 21, 2004 (received for review February 26, 2004)

The activation of mammalian origins of replication depends so far on ill understood epigenetic events, such as binding of transcription factors, chromatin structure, and nuclear localization. Understanding these mechanisms is not only a scientific challenge but also represents a prerequisite for the rational design of nonviral episomal vectors for mammalian cells. In this paper, we demonstrate that a tetramer of a 155-bp minimal nuclear scaffold/matrix attached region DNA module linked to an upstream transcription unit is sufficient for replication and mitotic stability of a mammalian episome in the absence of selection. Fluorescence *in situ* hybridization analyses, crosslinking with *cis*-diammineplatinum(II)-dichloride and chromatin immunoprecipitation demonstrate that this vector associates with the nuclear matrix or scaffold *in vivo* by means of specific interaction of the nuclear scaffold/matrix attached region with the nuclear matrix protein SAF-A. Results presented in this paper define the minimal requirements of an episomal vector for mammalian cells on the molecular level.

DNA replication | mitotic stability | nuclear matrix | SAF-A

In the budding yeast *Saccharomyces cerevisiae*, replication origins are defined entities that comprise a highly conserved AT-rich "core" consensus sequence. Although the activation of individual replication origins may depend on epigenetic factors (1, 2), all of these autonomously replicating sequences (ARSs) allow extrachromosomal replication when inserted into plasmids (3). In the fission yeast *Saccharomyces pombe*, replication origins are more extended than in budding yeast, and, although they show no consensus sequence in the strict sense, they contain AT-rich elements that serve as binding sites for the origin recognition complex (4). Metazoan origins of replication also contain restricted tracts of AT-rich sequences, but apart from this, they seem to be composed of discrete functional elements that differ among individual origins (5–8). The search for mammalian ARS-like elements was usually not successful, although a potential loose consensus sequence has been described for DNA circles that replicate episomally in mammalian cells (9). Still, the ultimate fate of all these constructs was either gradual loss during subsequent cell divisions or integration into the cellular genome. Epstein–Barr virus systems that entirely depend on the presence of Epstein–Barr virus-encoded nuclear antigen 1 have been described to promote long-term episomal replication (10).

It is now generally assumed that mammalian origins of replication are not exclusively determined by their DNA sequence, but their function relies on epigenetic principles, such as the presence of bound transcription factors, chromatin structure, or nuclear localization (11–13).

Understanding the control of replication is one of the most relevant problems to be solved in cell and molecular biology. This problem is not purely esoteric because this knowledge is a prerequisite for the construction of vectors that allow safe and reproducible genetic modification of cells and organisms (14). For both cases, an assay system comparable to the ARS system in yeast would be important. Recently, a prototype of such a system was designed in our laboratory (15). Based on the observation that the binding of an origin of replication to the nuclear matrix precedes the onset of S phase (12), the simian virus 40 (SV40), large T antigen was replaced by a nuclear scaffold/matrix attached region (S/MAR) thought to recruit the respective endogenous replication factors of the host cell. This vector, pEPI-1, was shown to replicate episomally at copy numbers of ≈5–10 in Chinese hamster ovary (CHO) (15), HeLa (16) and other cell lines, and it is mitotically stable over hundreds of generations in the absence of selection. It has been suggested that these properties can be explained by a specific interaction with the nuclear matrix by means of the constituent protein SAF-A (17, 18). Additionally, an active transcription unit upstream from the S/MAR is required to maintain the episomal status of the vector (19). However, the minimal functional elements required for episomal replication and mitotic stability have still to be defined, and we only now begin to understand the functioning of this vector type.

Here we report that a minimal S/MAR, i.e., a tetramer of an 155-bp module that has been identified in a natural S/MAR (20), can functionally replace the much more complex and extended original element. This sequence linked to an upstream transcription unit is the only requirement for the functioning of this vector as a mammalian episome. This construct represents a synthetic episomally replicating vector for mammalian cells suitable as an assay system for the study of replication control, which can also be seen as a vector prototype for a safe and reproducible genetic modification of mammalian cells and long-term expression of a vector-encoded transgene (21).

Methods

Cells, Vectors, and Transfection Procedure. CHO cells were grown under conditions described in ref. 15. Vectors used were pEPI-enhanced GFP (-eGFP), pGFP-C1, pDiMAR, pTetMAR, pEPI-Rous sarcoma virus (RSV), and pMARS (Fig. 1). The vector pEPI-RSV was constructed by deleting the SV40 origin/promoter sequence by digestion with *Sex*AI and *Stu*I and insertion of a *Sex*AI/*Stu*I fragment containing the RSV promoter. For the construction of pDiMAR and pTetMAR the dimer and tetramer of the S/MAR module was inserted into the vector

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: S/MAR, nuclear scaffold/matrix attached region; ARS, autonomously replicating sequences; SV40, simian virus 40; RSV, Rous sarcoma virus; CMV, cytomegalovirus, eGFP, enhanced GFP; FISH, fluorescence *in situ* hybridization; HAP, hydroxyapatite; CHO, Chinese hamster ovary.

†A.C.W.J. and I.M.S. contributed equally to this work.

¶To whom correspondence should be addressed at: Institute of Cell Biology, Witten/Herdecke University, Stockumer Strasse 10, D-58448 Witten, Germany. E-mail: lipps@uni-wh.de.

www.pnas.org/cgi/doi/10.1073/pnas.0401355101

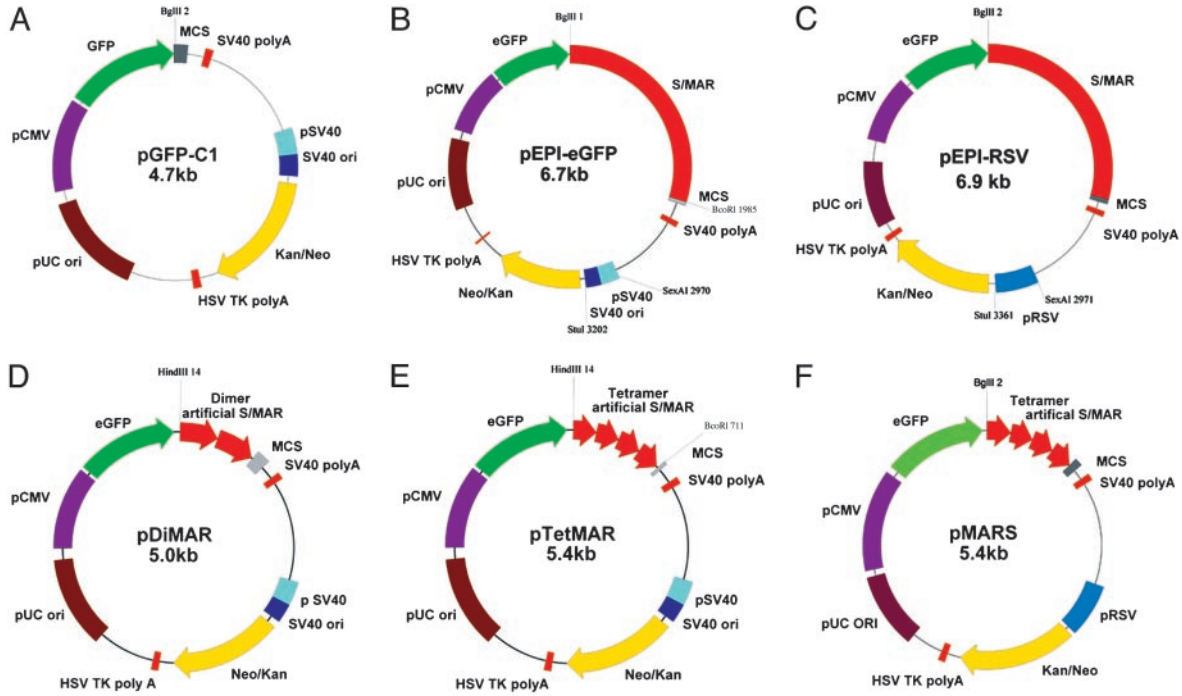

Fig. 1. Vectors used in this study are pGFP-C1 (Clontech) (*A*), pEPI-eGFP (*B*) (16), pEPI-RSV (*C*), pDiMAR (*D*), pTetMAR (*E*), and pMARS (*F*). For vector construction, see *Methods*. The regions are color-coded as follows. Pink box, CMV promoter; green arrow, GFP and eGFP; red box, S/MAR; gray box, multiple cloning site; orange box, SV40 polyadenylation sequence; blue box, SV40 origin of replication/promoter; yellow arrow, kanamycin/G418 resistance gene; orange box, herpes simplex virus thymidine kinase polyadenylation sequence; dark-red box, pUC origin of replication.

pTZ 18R (Sigma–Aldrich). From this vector, pDiMAR and pTetMAR could be isolated as an *Eco*RI/*Hin*dIII fragment and cloned into the multiple cloning site of pGFP-C1 (Clontech). The vector pMARS was assembled by excision of the S/MAR in pEPI-RSV and replacement by the *Bsp*EI/*Sma*I tetramer S/MAR module fragment of pTetMAR.

CHO cells were transfected with the various vectors by electroporation as described earlier (15). Two days after gene delivery, transfected cells were selected with 500 μg/ml G418 (Boehringer Mannheim, which is now Roche Molecular Biochemicals). Stable clones were isolated after 2–3 weeks and further cultured with or without selection pressure.

Isolation of DNA and Southern Blot Analysis. For Southern analysis (22) total cellular DNA and extrachromosomal DNA from a Hirt extract (23) from 10^7 cells were isolated, digested with the appropriate restriction enzyme, separated on 0.8% agarose gels, and blotted onto nylon membranes. The appropriate vector was labeled with ^{32}P (Ready-to-Go labeling kit, Amersham Pharmacia) and used as a probe. The hybridization was done in Church buffer (0.25 M sodium phosphate buffer, pH 7.2/1 mM EDTA/1% BSA/7% SDS) at 65°C for 16 h.

Rescue Experiments. Transformation of *E. coli* NovaBlue (Novagen–Calbiochem) with DNA of a Hirt extraction from 10^7 stably transfected CHO cells was performed by heat shock. *E. coli* transformants were selected by using agarose plates containing 30 μg/ml kanamycin. DNA was isolated from individual resistant clones and subjected to restriction analysis.

Isolation of RNA and Northern Blot Analysis. For Northern analysis (24) total RNA was isolated (25) from 10^6 cells, separated on 1.2% formaldehyde–agarose gels, transferred to nylon membranes, and hybridized with a ^{32}P-labeled eGFP gene as a probe.

Crosslinking with *cis*-Diammineplatinum(II)-dichloride. Cells ($n = 4 \times 10^7$) were collected by centrifugation, and the *cis*-platin crosslinking was performed as described by Ferraro and coworkers (26, 27). Crosslinking was stopped, and the cells were resuspended in 10 ml of lysis buffer (5 M urea/2 M guanidine hydrochloride/2 M NaCl/1 mM PMSF) and bound to hydroxyapatite (HAP, Bio-Rad). Desorption of DNA and DNA/protein complexes from HAP and their concentration were performed as described in ref. 18, but desorbed material was concentrated by precipitation with ethanol.

The presence of vector DNA was detected by PCR analysis by using 250 ng of DNA as a template. Primers used for PCR analysis were derived from the cytomegalovirus (CMV) promoter and had the sequences CMV43–62 5′-CCC ATA TAT GGA GTT CCG-3′ and CMV795–776 5′-GAC AAG TGT TGG CCA TGG AA-3′. PCR was performed as described in refs. 17 and 18.

***In Vivo* Crosslinking with Formaldehyde.** Formaldehyde was diluted to 1% in prewarmed culture medium (37°C) and added to monolayers of 10^8 cells for 10 min. Cells were washed three times with cold PBS, scraped off, washed twice with PBS, and then resuspended in hypotonic RSB buffer (10 mM Tris·HCl/3 mM $MgCl_2$/10 mM sodium bisulfite, pH 8.0). After 10 min on ice, the swollen cells were disrupted by Dounce homogenization. The nuclear material was collected and washed twice with ice cold RSB buffer, once with high-salt NSB buffer (1 M NaCl/10 mM Tris·HCl/0.1% Nonidet P-40/1 mM EDTA, pH 7.5), and once with low-salt NSB (0.1 M NaCl). The material was then resuspended in TE/sodium bisulfite (10 mM Tris·HCl/1 mM EDTA/10 mM sodium bisulfite, pH 7.4) and sonicated briefly on ice. The probes were adjusted to an OD_{260} of 40.

Chromatin Immunoprecipitation. Chromatin immunoprecipitations were performed in NET buffer (50 mM Tris·HCl, pH

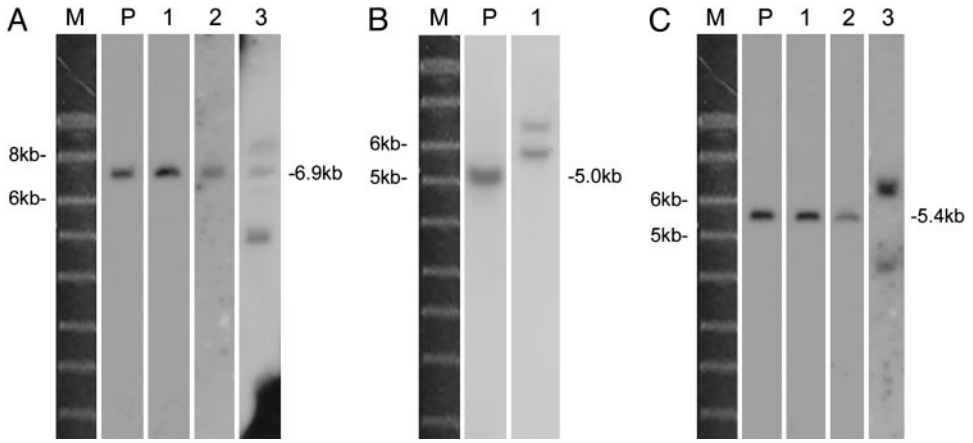

Fig. 2. Southern analyses of DNA isolated from CHO cells transfected with pEPI-RSV (*A*), pDiMAR (*B*), and pTetMAR (*C*). Hybridization was done as described in *Methods*. The hybridization pattern of one representative clone is shown for each construct. Lanes: M, DNA Marker *SMART-Ladder* (Eurogentec, Brussels); P, pEPI-RSV (*A*), pDiMAR (*B*), or pTetMAR (*C*) plasmid DNA as a control linearized by digestion with *Bgl*II (*A*) or *Hin*dIII (*B* and *C*); 1, total DNA digested with *Bgl*II (*A*) or *Hin*dIII (*B* and *C*); 2, Hirt extract digested with *Bgl*II (*A*) or HindIII (*C*); 3, undigested DNA from a Hirt extract.

7.4/150 mM NaCl/5 mM EDTA/0.5% Nonidet P-40). Nucleoprotein probe (440 μl) was incubated with 10 μl of anti-SAF-A antibody (28) for 12 h at 4°C. Immunocomplexes were collected by centrifugation (10,000 $\times$ *g*, 2 h, 4°C), and DNA or proteins were isolated.

DNA (250 ng) was subjected to PCR analysis and protein was subjected to a Western analysis by using an anti-SAF-A antibody (28). Primers used for PCR analysis were derived from the neomycin gene and had the sequences Neo9left 5′-GCTTGCCGAATATCATGGTG-3′ and Neo9right 5′-GGGTTCCTTCCGGTATTGTC-3′. PCR and radioactive PCR were performed as described in refs. 17 and 18 and in the supporting information, which is published on the PNAS web site. The products were analyzed after 0, 15, 20, 25, 30, 35, and 50 cycles. As internal control, the minimal S/MAR element of the CHO *dhfr*-locus was amplified from the immunoprecipitates and subjected to the same PCR analyses by using the primers Mesner_left 5′-AGCGGTCCACCAAACTTTTT-3′ and Mesner_right 5′-AGTCACCATTTATAGGGGAGAAT-3′ (29).

***In Vitro* Binding Assays.** *In vitro* binding assays to the isolated DNA-binding domain of SAF-A were carried out as described in ref. 28. As a substrate, the S/MAR module dimer or tetramer was amplified and radiolabeled by PCR with a primer pair specific for vector sequences flanking the insert.

Chromosome Spreading and Fluorescence *in Situ* Hybridization (FISH) Analysis. FISH analyses were done by using biotin-labeled pGFP-C1 as a probe. After hybridization in 50% formamid/10% dextransulfat/2× SSC/40 mM $NaHPO_4$/SDS 0.1%/1× Denhardt's solution, pH 7.0, at 37°C for 16 h, the labeled probe was detected by an avidin/tetramethyl rhodamine isothiocyanate-labeled anti-avidin antibody/biotin (Sigma) sandwich procedure according to Liu *et al.* (30). After immunostaining, the metaphase chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (Sigma-Aldrich) and analyzed with a Leica DM RB fluorescence microscope equipped with a Leica DC 300f camera.

Results

Vectors containing only a S/MAR sequence or the SV40 origin of replication integrate into the genome of the recipient CHO cell. It was originally concluded that the functioning of the pEPI-1 episome relies on the combination of the SV40 origin linked to a chromosomal S/MAR (15). More recently it became obvious that an upstream transcription running into the S/MAR is another essential feature required for the episomal status. Deletion of this transcription unit or reversing its direction leads to vector integration (19). In addition, it was shown that the origin recognition complex and Mcm proteins can assemble at random sites on this vector and DNA replication initiation can occur at various sites (16). Based on these observations we analyzed whether episomal replication of pEPI-1 can still be observed after deleting the SV40 origin of replication. For this purpose, the SV40 origin was deleted in construct pEPI-RSV (Fig. 1C) and CHO cells were transfected by electroporation with the supercoiled plasmid and selected with G418 (15). Either total DNA or DNA from the Hirt extract was isolated from 30 clones and subjected to Southern analysis. The pattern of DNA was found to be identical to the original vector DNA in 28 of 30 clones (one example is shown in Fig. 2*A*). Only in two clones the plasmid became integrated. These data are very similar to those obtained with the original vector pEPI-1 where the percentage of episomally replicating entities depends on the superhelical status of input vector DNA, i.e., any nicked circle or linearized vector DNA becomes subject to integration (15, 19). Based on the intensity of the hybridization signal, the copy number of the episomal vector was estimated to be $<$20 as previously reported for pEPI-1 (15–17). These results strongly indicate that it is only the S/MAR linked to an upstream transcription that is required for episomal replication in CHO cells.

In a parallel set of experiments, we defined the minimal S/MAR DNA element that would sustain episomal replication and mitotic stability. S/MARs represent base-unpairing regions, which, according to certain requirements (unwinding element spacing and overall length) associate with specialized proteins (SAF-A and lamins) and support dynamic association of any linked gene with the nuclear matrix (20). By using these criteria, a 155-bp S/MAR module with a 25-bp linker for oligomerization was designed that comprises the core unwinding element of the human IFN upstream S/MARs (31). Based on the observation that only oligomers of the S/MAR module can augment transcription in a standard assay (for details, see supporting information) (32), we decided to replace the 2-kb chromosomal S/MAR by either the dimer (pDiMAR) or tetramer (pTetMAR) of this 180-bp sequence. Of the clones of CHO cells transfected by pDiMAR (Fig. 1*D*) and pTetMAR (Fig. 1*E*), 15 were isolated and their DNA subjected to Southern analysis. Although all vectors containing the dimer S/MAR module integrated into the genome (Fig. 2*B*), in all 15 clones containing the tetramer S/MAR module, the pattern of DNA was identical to that

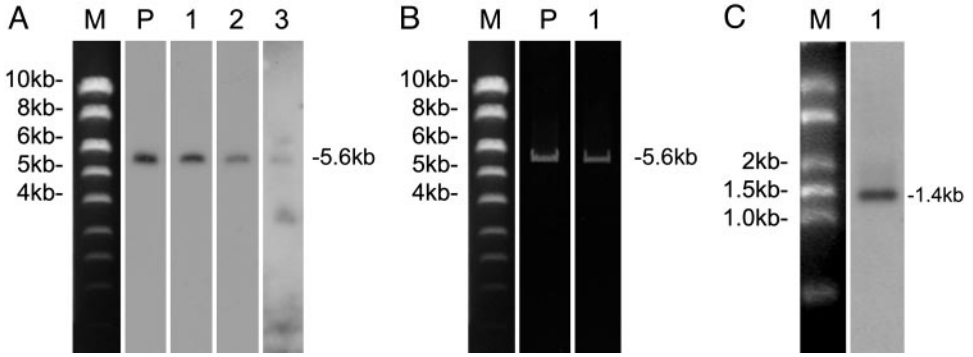

Fig. 3. Binding of pMARS in the cell. (*A*) Southern analyses of DNA isolated from CHO cells transfected with pMARS. Hybridization was done as described in *Methods*. The hybridization pattern of one representative clone is shown. (*B*) Rescue experiment in *E. coli* with Hirt extract from CHO cells transfected with pMARS. (*C*) Northern analysis of total RNA isolated from CHO cells transfected with pMARS. Hybridization was done as described in *Methods*. The hybridization pattern of one representative clone is shown. Lanes: M, DNA Marker *SMART-Ladder* (Eurogentec); P, pMARS (400 ng in *B*) plasmid DNA as a control linearized by digestion with *Hin*dIII; 1, total DNA digested with *Hin*dIII (*A*), DNA isolated from one bacterial clone digested with *Hin*dIII (*B*), or total RNA hybridized with a ^{32}P-labeled eGFP probe (*C*); 2, Hirt extract digested with *Hin*dIII; 3, undigested DNA from a Hirt extract.

of the input vector and the vectors supercoiled form could be isolated from the Hirt extract (Fig. 2*C*). Dimers of the amplification-promoting sequences NTS-1 and NTS-2 from the mouse rRNA-encoding DNA cluster (33), which have a similar AT content and length but lack the extended base unpairing region (17) of the tetramer S/MAR module, did not support episomal replication.

These results suggest that it is possible to construct an episomally replicating vector whose function relies only on the tetramer of a S/MAR module and an upstream transcription unit. This vector, pMARS, shown in Fig. 1*F*, was transfected into CHO cells and the DNA of 10 G418-resistant clones analyzed. The vector restriction pattern of all clones was identical to that of the original vector, and the supercoiled form of this vector could be isolated by Hirt extraction. One example of these analyses is shown in Fig. 3*A*. To further verify the episomal nature of the vector, DNA from Hirt extraction was used to transfect *E. coli*. In none of the 12 analyzed *E. coli* transfectants was DNA rearrangement detected in the rescued vector. Fig. 3*B* shows the restriction pattern of one transfected *E. coli* clone. The mitotic stability of the vector pMARS was assessed by growing transfected CHO cells in the absence of selection for a period of 8 weeks followed by selection with G418. It could be calculated that the mitotic stability of this construct was >0.98, a number similar to that observed for pEPI-1-transfected HeLa cells (16). This mitotic stability became also evident when DNA was isolated from a transfected clone after 8 weeks of growth in nonselective medium that showed no change in hybridization intensity relative to DNA isolated from the same clone immediately after selection (data not shown).

We have reported earlier (19) that ongoing transcription upstream of the S/MAR is required for episomal replication and that the transcript has to include at least parts of the S/MAR. By using the eGFP gene as a probe, a Northern analysis was performed with total RNA from a pMARS-transfected (Fig. 1*F*) clone. As shown in Fig. 3*C*, an $\approx$1.4-kb transcript can be found, suggesting that transcription traverses the entire tetrameric element, terminating at the SV40 polyadenylation signal. Deletion of this transcription unit or placing the SV40 polyadenylation site between eGFP and S/MAR resulted in integration of the vector (data not shown).

It has been described that pEPI-1 binds to the nuclear matrix *in vivo* through a specific interaction of the S/MAR with the matrix protein SAF-A (18). We therefore performed FISH analysis on spread CHO chromosomes transfected either with the integrating pDiMAR vector DNA (Fig. 1*E*) or the episomally replicating pMARS vector DNA (Fig. 1*F*). About 50 metaphase plates were analyzed for each of these two FISH analyses. Although pDiMAR DNA is always found integrated at a single chromatid, in all clones different site and hybridization to both chromatids is observed (Fig. 4*A Left*), the numbers of signals obtained with pMARS DNA varied in different cells of the same clone and were in most cases associated with only one chromatid. Depending on the shear forces applied during spreading free molecules could also be found (Fig. 4*A Right*). The average copy number estimated from this analysis was found to be 6, ranging between 4 and 11 per cell (for further informa-

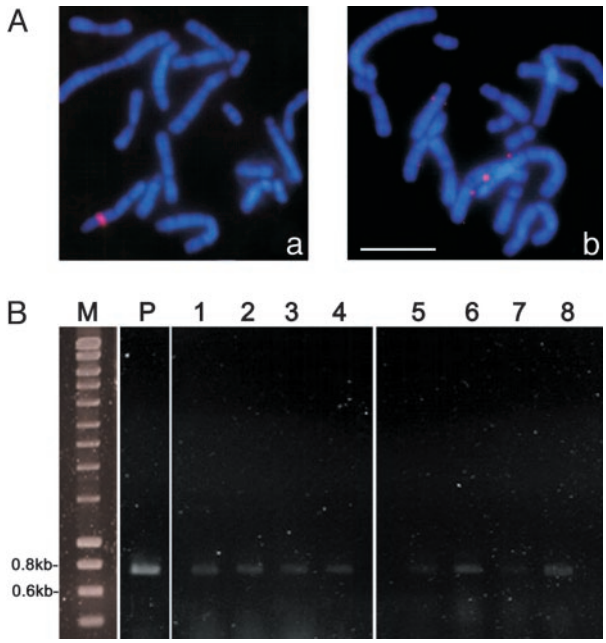

Fig. 4. Binding of pMARS to the nuclear matrix. (*A*) FISH analyses of CHO cells transfected either with pDiMAR (*Left*) or with pMARS (*Right*) were done as described in *Methods*. (Scale bar, 10 μm.) (*B*) PCR analyses from DNA bound to HAP after crosslinking of transfected CHO cells by using 250 ng of DNA as a template. Lanes: 1, supernatant from pGFP-C1-transfected cells; 2, supernatant from pEPI-eGFP-transfected cells; 3, supernatant from pDiMAR-transfected cells; 4, supernatant from pMARS-transfected cells; 5, DNA bound to HAP from pGFP-C1-transfected cells; 6, DNA bound to HAP from pEPI-eGFP-transfected cells; 7, DNA bound to HAP from pDiMAR-transfected cells; 8, DNA bound to HAP from pMARS-transfected cells; 9, control PCR with 100 pg of pEPI-eGFP vector as a template.

CELL BIOLOGY

tion, see supporting information), which is very similar to that reported for pEPI-1 in CHO (17) and HeLa cells (16).

To further demonstrate the association of vector pMARS with the nuclear matrix, *in vivo* crosslinking experiments were performed by using *cis*-diammineplatinum(II)-dichloride, an agent that links matrix proteins to S/MAR DNA with high specificity (26). Cells transfected with the integrating vectors pGFP-C1 and pDiMAR and the episomally replicating vector pEPI-eGFP and pMARS (Fig. 1) were subjected to this analysis. Cells were crosslinked with 1 mM *cis*-diammineplatinum(II)-dichloride and lysed in the appropriate buffer (see *Methods*). The DNA was then fragmented by sonification, and the matrix-associated DNA–protein complex was adsorbed by HAP, whereas all noncrosslinked material stayed in the supernatant (27). DNA from the supernatant and the HAP-bound material was isolated, and identical amounts of DNA were used for PCR analysis by using primers derived from the CMV promoter (17, 18). As shown in Fig. 4*B*, vector DNA from the HAP-bound material could only be amplified from cells transfected by pEPI-eGFP and pMARS; some vector DNA also occurred in the supernatant fraction. In contrast, for the pGFP-C1 and pDiMAR vector, no or only a minor signal could be obtained from the HAP-bound material, but a strong signal was consistently obtained in the supernatant fraction.

Because pEPI-1 was shown to associate with the nuclear matrix by means of specific interaction with SAF-A (18), we investigated whether the episomal status of the vectors containing the S/MAR tetramer module (pMARS and pTetMAR) but not pDiMAR (containing the S/MAR dimer module) could reflect differences in the binding to this protein. Indeed, by using the well characterized pull-down assay of Kipp *et al.* (28), we found that only the tetrameric, and not the dimeric, S/MAR module was specifically recognized by the isolated DNA-binding domain of SAF-A (Fig. 5*A*). This result reflects the mass-binding mechanism of SAF-A (28), in which multiple weak interactions sum up to highly specific binding when the DNA element is long enough. Consistent with these *in vitro* binding data, chromatin immunoprecipitation experiments with SAF-A antibodies confirm that pTetMAR (and pEPI-eGFP as a positive control), but not pDiMAR or the control vector pGFP-C1, are bound to SAF-A in living cells (Fig. 5*B*). Various clones transfected with pGFP-C1, pDiMAR, pTetMAR, and pEPI-eGFP were *in vivo* crosslinked with formaldehyde and precipitated with an antibody directed against SAF-A (34). Vector DNA in the precipitate from the different clones was characterized by PCR analysis by using primers derived from the neomycin-resistance cassette. Equal aliquots of the precipitate were subjected to Western analysis by using a SAF-A-specific antibody, verifying that the amount of precipitated SAF-A protein was identical in all samples (Fig. 5*B Bottom*). An endogenous S/MAR element from the host cell, the minimal S/MAR of the *dhfr* locus (29), was also present in all immunoprecipitates in identical amounts, demonstrating equal amounts of amplifiable DNA in all samples (Fig. 5*B Middle*). In striking contrast, however, the amount of vector-specific PCR products varied dramatically between the plasmids under study. Although the amount of the PCR products from pEPI-eGFP and pTetMAR where almost identical on an ethidium bromide-stained gel, only weak bands were obtained from pDiMAR and pGFP-C1-transfected cells (Fig. 5*B Top*). To quantify this difference, a radioactive PCR was performed with immunoprecipitates from three independent clones, and the products were analyzed after various PCR cycles (for details, see supporting information). By using the PCR product from pEPI-eGFP-transfected cells as a reference (100%), the concentration of the PCR products from pTetMAR-transfected cells varied between 78% and 99%. In contrast, the concentration of PCR products obtained from pDiMAR- and pGFP-C1-transfected cells varied between 7% and 11% and between 5% and 6%, respectively. By using unspecific antibodies for immunoprecipitation (anti-Oxytricha βTP and anti-rabbit IgG) in no case vector DNA could be amplified from the precipitate.

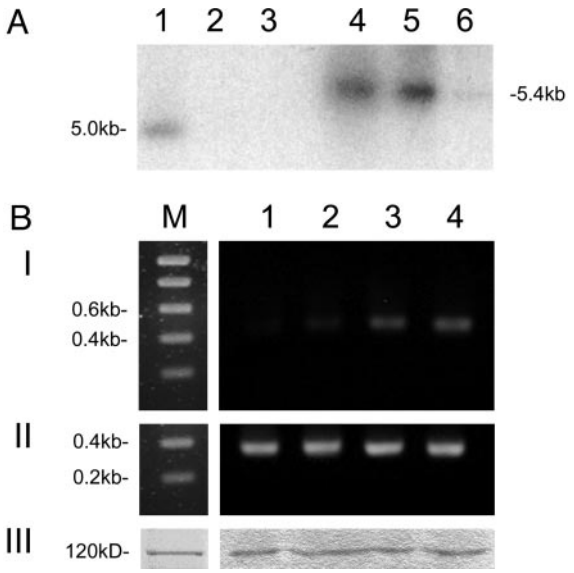

Fig. 5. Binding of pMARS to SAF-A. (*A*) *In vitro* pull-down assay with pDiMAR as a template and without competitor DNA (lane 1), with a 100-fold excess of *E. coli* competitor DNA (lane 2), and with a 10,000-fold excess of *E. coli* competitor DNA (lane 3) or with pTetMAR as a template and without competitor DNA (lane 4), with a 100-fold excess of *E. coli* competitor DNA (lane 5), and with a 10,000-fold excess of *E. coli* competitor DNA (lane 6). (*B*) PCR analyses from DNA (*Top*) and Western analyses (*Bottom*) with an anti-SAF-A antibody from proteins isolated after *in vivo* crosslinking of transfected CHO cells and immunoprecipitation. As an internal control, the endogenous minimal S/MAR of the *dhfr*-locus (29) was amplified (30 cycles) from the immunoprecipitates (*Middle*). (*Top*) PCR analysis of DNA isolated from an immunoprecipitate of pGFP-C1- (lane 1), pDiMAR- (lane 2), pTetMAR- (lane 3), or pEPI-eGFP-transfected (lane 4) cells. (*Middle*) Amplification of the minimal S/MAR of the *dhfr* locus. (*Bottom*) Western analysis of immunoprecipitates by using an anti-SAF-A antibody. Lanes correspond to *Top*.

Discussion

In this report we define the requirements for persistence and replication of a mammalian episome exclusively based on mammalian chromosomal elements. This task was achieved by modification of the previously described vector pEPI-1 in which the role of the SV40 large T-antigen is taken over by a chromosomal S/MAR (15). In a parallel set of experiments we showed that the SV40 origin of the previously described vector pEPI-1 is not required for replication of a mammalian episome, and we defined the minimal sequence requirements for a S/MAR to function *in vivo*.

Southern analyses of constructs from which the SV40 origin was deleted showed that this sequence is not required for vector function: in >90% of transfected cells, its restriction pattern was identical to that of the input DNA. When the 2-kb chromosomal S/MAR was replaced by oligomers of a 155-bp minimal S/MAR module, vectors containing the dimer integrate into the genome, whereas vectors containing a tetramer of this sequence remain episomal and yield a restriction pattern identical to that of the input vector. Other AT-rich sequences of similar length did not support episomal replication, demonstrating that this ability is due to the structural characteristics of the S/MAR module (17, 32). Based on these results, we designed an episomal vector whose function relies exclusively on an interaction between an active transcription unit and a minimal S/MAR module.

The resulting vector, pMARS, contains a tetramer of the S/MAR module with an upstream expression cassette. As demonstrated by Southern analyses and rescue experiments, pMARS replicates episomally in CHO cells and is mitotically stable in the absence of selection. Transcription from an upstream expression cassette has to include the S/MAR because placing a termination site between the gene and the S/MAR results in integration. The association of this minimal vector with the nuclear matrix was demonstrated both by FISH analyses and by *in vivo* crosslinking with *cis*-diammineplatinum(II)-dichloride. In FISH analyses, an association of the vector with the mitotic chromosome could be shown and its noncovalent nature could be demonstrated by applying strong shear forces during spreading. Such a treatment dissociates the vector from the chromatids as already described for pEPI-1 (17).

In nuclear matrix preparations only the vectors pEPI-eGFP and pMARS copurify with the matrix fraction, although some nonbound vector material is also found in the supernatant. This phenomenon may be due to incomplete binding of the S/MAR–protein complex to HAP as frequently observed with this technique (27). Both *in vitro* and *in vivo*, only the tetrameric, but not the dimeric, S/MAR construct is recognized by the nuclear matrix protein SAF-A, as shown for pEPI-1 (18). The difference between the two elements reflects the peculiar DNA-binding mode of SAF-A, which recognizes S/MAR DNA by means of mass-binding (28). In this binding mode, short AT-rich sequences are weakly bound by individual DNA-binding domains of SAF-A, but the clustering of such sequences on a contiguous DNA strand sum up to highly specific binding to self-assembled SAF-A protein complexes (28, 34). This explains why multimers of identical sequence elements are only recognized by SAF-A when they exceed a minimal length, a prerequisite that is obviously fulfilled by the tetrameric, but not the dimeric, S/MAR module construct.

Results presented in this study now allow the proposal of a mechanistic explanation for the episomal replication and maintenance of a vector in mammalian cells. The presence of a S/MAR DNA element tethers the vector to the nuclear matrix and to its derivative during mitosis, explaining the mitotic stability by piggybacking the vector on chromosomes during nuclear division. This suggestion is consistent with a recent report showing that S/MARs might be involved in cohesion and separation of chromatids (29). This association allows the vector to use the cellular replication machinery, explaining why it is replicated only once in early S phase during the cell cycle (16). Therefore, the S/MAR sequence contained in pMARS may have a similar function as the Epstein–Barr virus-encoded nuclear antigen protein in Epstein–Barr virus-derived vectors. In addition, a transcription running into or through the S/MAR appears to be necessary, most probably because it creates a chromatin structure that is accessible for replication enzymes.

These results together with earlier observations strongly support the idea that a mammalian origin of replication is not simply defined by a specific DNA sequence but by epigenetic characteristics, such as chromatin structure, nuclear localization, binding of transcription factors, an interplay between transcription and replication, and the ability to bind the origin recognition complex (16, 19, 35). The vector pMARS contains only the minimal functional elements required for correct replication and mitotic stability and can be manipulated with ease, thus providing an excellent model system for the analysis of the epigenetic control of DNA replication in mammalian cells.

This work was supported by the Deutsche Forschungsgemeinschaft, the European Union, and the Alfried Krupp von Bohlen und Halbach Foundation. A.C.W.J. is the recipient of a postdoctoral fellowship from the Boehringer Ingelheim Stiftung.

1. Muller, M., Lucchini, R. & Sogo, J. M. (2000) *Mol. Cell* **5,** 767–777.
2. Pasero, P. & Schwob, E. (2000) *Curr. Opin. Genet. Dev.* **10,** 178–186.
3. Stinchcomb, D. T., Thomas, M., Kelly J., Selker, E. & Davis, R. W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4559–4563.
4. Chuang, R. Y. & Kelly, T. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 2656–2661.
5. Altman, A. L. & Fanning, E. (2001) *Mol. Cell. Biol.* **21,** 1098–1110.
6. DePamhilis, M. L. (1999) *BioEssays* **21,** 186–197.
7. Gilbert, D. M. (2001) *Science* **294,** 96–100.
8. Liu, G., Malott, M. & Leffak, M. (2003) *Mol. Cell. Biol.* **23,** 1832–1842.
9. Price, G. B., Allarakhia, M., Cossons, N., Nielsen, T., Diaz-Perez, M., Friedlander, P., Tao, L. & Zannis-Hadjopoulos, M. (2003) *J. Biol. Chem.* **278,** 19649–19659.
10. Krysan, P. J., Haase, S. B. & Calos, M. P. (1989) *Mol. Cell. Biol.* **9,** 1026–1033.
11. Biamonti, G., Paixao, S., Montecucco, A., Peverali, F. A., Riva, S. & Falaschi, A. (2003) *Chromosome Res.* **11,** 403–412.
12. Cook, P. R. (1999) *Science* **284,** 1790–1795.
13. Mechali, M. (2001) *Nat. Rev. Genet.* **2,** 640–645.
14. Lipps, H., Jenke, A., Nehlsen, K., Scinteie, M., Stehle, I. M. & Bode, J. (2003) *Gene* **11,** 403–412.
15. Piechaczek, C., Fetzer, C., Baiker, A., Bode, J. & Lipps, H. J. (1999) *Nucleic Acids Res.* **27,** 426–428.
16. Schaarschmidt, D., Baltin, J., Stehle, I. M., Lipps, H. J. & Knippers, R. (2004) *EMBO J.* **23,** 191–201.
17. Baiker, A., Maercker, C., Piechaczek, C., Schmidt, S. B. A., Bode, J., Benham, C. & Lipps, H. J. (2000) *Nat. Cell Biol.* **2,** 182–184.
18. Jenke, B. H. C., Fetzer, C. P., Stehle, I. M., Jönsson, F., Fackelmayer, F. O., Conradt, H., Bode, J. & Lipps, H. J. (2002) *EMBO Rep.* **3,** 349–354.
19. Stehle, I. M., Scinteie, M. F., Baiker, A., Jenke, A. C. & Lipps, H. J. (2003) *Chromosome Res.* **11,** 413–421.
20. Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klohr, D., Mielke, C. & Kohwi-Shigematsu, T. (1992) *Science* **255,** 195–197.
21. Jenke, A. C., Sincteie, M., Stehle, I. M. & Lipps, H. J. (2004) *Mol. Biol. Rep.* **31,** 85–90.
22. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
23. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369.
24. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5350–5354.
25. Chomczynski, P. (1993) *BioTechniques* **15,** 532–537.
26. Eufemi, M., Ferraro, M., Altieri, F., Cervoni, L. & Turano, C. (2000) *Mol. Biol. Rep.* **27,** 181–189.
27. Ferraro, A., Cervoni, L., Eufemi, M., Altieri, F. & Turano, C. (1996) *J. Cell. Biochem.* **62,** 495–505.
28. Kipp, M., Gohring, F., Ostendorp, T., van Drunen, C. M., van Driel, R., Przybylski, M. & Fackelmayer, F. O. (2000) *Mol. Cell. Biol.* **20,** 7480–7489.
29. Mesner, L. D., Hamlin, J. L. & Dijkwel, P. A. (2003) *Proc. Natl. Acad. Sci. USA* **18,** 3281–3286.
30. Liu, C. G., Maercker, C., Castanon, M. J., Hauptmann, R. & Wiche, G. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 4278–4283.
31. Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T. & Bode, J. (1990) *Biochemistry* **29,** 7475–7485.
32. Benham, C., Kohwi-Shigematsu, T. & Bode, J. (1997) *J. Mol. Biol.* **272,** 181–196.
33. Wegner, M., Zastrow, G., Klavinius, A., Schwender, S., Muller, F., Luksza, H., Hoppe, J., Wienberg, J. & Grummt, F. (1989) *Nucl. Acids Res.* **17,** 9909–9932.
34. Fackelmayer, F. O. & Richter, A. (1994) *Biochemistry* **33,** 10416–10422.
35. Vashee, S., Cevtic, C., Lu, W., Simancek, P., Kelly, T. J. & Walter, C. (2003) *Genes Dev.* **17,** 1894–1908.